Mission ImPASsible:
Decoding Polycyclic Aromatic Systems with Deep Learning

Renana Gershoni-Poranne

Schulich Faculty of Chemistry, Technion – Israel Institute of Technology, Israel

Polycyclic aromatic systems (PASs) present a seemingly insurmountable challenge: vast chemical spaces, complex electronic structures, and elusive aromatic properties. Our mission, should we choose to accept it, is to harness the power of deep learning to decode these molecular mysteries. In this talk, we embark on a journey through this complex chemical space, combining traditional computational methods with cutting-edge artificial intelligence tools. We demonstrate how neural networks can be trained to predict electronic properties with unprecedented speed and accuracy. More importantly, we show how they can be used interpretably to extract chemical insight.

To enable the application of such techniques to the design of novel functional PASs, we established the COMPAS Project – a COMputational database of Polycyclic Aromatic Systems – which already contains ~1 million molecules in three datasets [1−4]. We also developed two new types of molecular representation to enable efficient and effective machine- and deep-learning models to train on the new data: a) a text-based representation [5] and b) a graph-based representation [6]. By analyzing thousands of PAS structures with our dedicated representations, our AI agents not only achieve higher predictive ability with fewer data but have also uncovered hidden patterns and structure-property relationships that traditional methods might have missed.

Finally, we implemented the first guided diffused-based model for inverse design of PASs: GaUDI [7]. Our model generates new PASs with defined target properties. In addition to its flexible target function and high validity scores, GaUDI also accomplishes design of molecules with properties beyond the distribution of the training data.

From small benzenoid systems to extended graphene-like structures, we show how deep learning can navigate this complex chemical landscape, offering new insights into molecular design and property prediction. This talk will not self-destruct in five seconds, but it will revolutionize how we think about combining artificial intelligence with molecular science.

[1] A. Wahab, L. Pfuderer, E. Paenurk, R. Gershoni-Poranne, *J. Chem. Inf. Model.* **2022**, *62*, 3704.
[2] E. Mayo Yanes, S. Chakraborty, R. Gershoni-Poranne, *Sci. Data* **2024**, *11*, 97.
[3] S. Chakraborty, E. M. Yanes, R. Gershoni-Poranne, *Beilstein J. Org. Chem.* **2024**, *20*, 1817–1830.
[4] A. Wahab, R. Gershoni-Poranne, *Phys. Chem. Chem. Phys.* **2024**, 26, 15344–15357.
[5] S. Fite, A. Wahab, E. Paenurk, Z. Gross, R. Gershoni-Poranne, *J. Phys. Org. Chem.* **2022**, e4458.
[6] T. Weiss, A. Wahab, A. M. Bronstein, R. Gershoni-Poranne, *J. Org. Chem.* **2023**, *88*, 9645–9656.
[7] T. Weiss, E. M. Yanes, S. Chakraborty, L. Cosmo, A. M. Bronstein, R. Gershoni-Poranne, *Nat. Comput. Sci.* **2023**, *3*, 873–882.